



Pre
UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/729,240	12/05/2000	Krishna Bharat	0026-0001	4725
26615	7590	05/28/2004	EXAMINER	
HARRITY & SNYDER, LLP 11240 WAPLES MILL ROAD SUITE 300 FAIRFAX, VA 22030			MAHMOUDI, HASSAN	
		ART UNIT		PAPER NUMBER
		2175		0
DATE MAILED: 05/28/2004				

Please find below and/or attached an Office communication concerning this application or proceeding.

P26

Office Action Summary	Application No.	Applicant(s)
	09/729,240	BHARAT ET AL.
	Examiner	Art Unit
	Tony Mahmoudi	2175

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

- 1) Responsive to communication(s) filed on 23 March 2004.
- 2a) This action is FINAL. 2b) This action is non-final.
- 3) Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

- 4) Claim(s) 1-41 is/are pending in the application.
- 4a) Of the above claim(s) _____ is/are withdrawn from consideration.
- 5) Claim(s) _____ is/are allowed.
- 6) Claim(s) 1-41 is/are rejected.
- 7) Claim(s) _____ is/are objected to.
- 8) Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

- 9) The specification is objected to by the Examiner.
- 10) The drawing(s) filed on _____ is/are: a) accepted or b) objected to by the Examiner.
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

- 12) Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) All b) Some * c) None of:
 1. Certified copies of the priority documents have been received.
 2. Certified copies of the priority documents have been received in Application No. _____.
 3. Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

SUPERVISORY PATENT EXAMINER
TECHNOLOGY CENTER 2100

Attachment(s)

- | | |
|-------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------|
| 1) <input type="checkbox"/> Notice of References Cited (PTO-892) | 4) <input type="checkbox"/> Interview Summary (PTO-413) |
| 2) <input type="checkbox"/> Notice of Draftsperson's Patent Drawing Review (PTO-948) | Paper No(s)/Mail Date. _____ |
| 3) <input type="checkbox"/> Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08)
Paper No(s)/Mail Date _____. | 5) <input type="checkbox"/> Notice of Informal Patent Application (PTO-152) |
| | 6) <input type="checkbox"/> Other: _____. |

DETAILED ACTION

Remarks

1. In response to communications filed on 23-March-2003, claims 1-41 are presently pending in the application.

Claim Rejections - 35 USC § 103

2. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

3. Claims 1-3, 5-8, 10-15, 17-22, 24-27, 29-32, 34, and 36-41 are rejected under 35 U.S.C. 103(a) as being unpatentable over Nanjo et al (U.S. patent No. 5,778,361) in view of Pant et al (U.S. patent No. 6,012,053.)

As to claim 1, Nanjo et al teaches a method of identifying (see Abstract) semantic units within a search query (see column 5, lines 23-32, where “semantic units” is read on “text in compound-word languages”, and see column 8, lines 21-33, where “semantic units” is read on “combination of text strings”) comprising:

identifying documents relating to the query (see column 6, lines 51-58) by comparing search terms in the query to an index of a corpus (see Abstract, and see column 21, lines 29-30);

generating a plurality of multiword substrings from the query (see column 4, lines 4-12, and see column 8, lines 21-33, where "multiword substrings" is read on "combination of text strings") in which each of the substrings includes at least two words (see column 8, line 33).

It is inherent that a "phrase" contains at least two words");

Nanjo et al does not teach calculating, for each of the generated substrings, a value that corresponds to a comparison between one or more of the identified documents and the generated substrings; and selecting semantic units from the generated multiword substrings based on the calculated values.

Pant et al teaches a system for relevance-ranking of search results (see Abstract), in which he teaches calculating, for each of the generated substrings, a value that corresponds to a comparison between one or more of the identified documents and the generated substrings (see column 7, lines 7-50); and selecting semantic units from the generated multiword substrings based on the calculated values (see column 10, lines 38-49, and see column 13, lines 9-25.)

Therefore, it would have been obvious to a person having ordinary skill in the art at the time the invention was made to have modified Nanjo et al as modified to include calculating, for each of the generated substrings, a value that corresponds to a comparison between one or more of the identified documents and the generated substrings; and selecting semantic units from the generated multiword substrings based on the calculated values.

It would have been obvious to a person having ordinary skill in the art at the time the invention was made to have modified Nanjo et al as modified, by the teaching of Pant et al, because calculating, for each of the generated substrings, a value that corresponds to a

comparison between one or more of the identified documents and the generated substrings; and selecting semantic units from the generated multiword substrings based on the calculated values, would enable the system to define and apply relevancy values (factors) between documents as well as between contents of documents in order to identify the most relevant (strongest) match between the search terms and the identified terms within a query.

As to claims 2 and 26, Nanjo et al as modified teaches wherein the identification of the documents further includes:

generating an initial list of relevant documents (see Nanjo et al, column 3, lines 19-26, where “list of relevant documents” is read on “list of files or documents that satisfy the search criteria”); and

selecting a predetermined number of most relevant ones of the documents (see Nanjo et al, column 17, lines 35-50, where “most relevant” is read on “exact match”) in the initial list as the identified documents (see Nanjo et al, column 8, lines 1-5, where “selecting documents” is read on “editing a particular document”).

As to claims 3, 8, 15, 22, 27, and 32, Nanjo et al as modified teaches wherein the selection of the semantic units further includes:

selecting semantic units from the generated substrings that have calculated values above a predetermined threshold (see Nanjo et al, column 20, lines 41-50, where “predetermined threshold” is read on “predetermined step size”, also see Pant et al, column 8, lines 57-62.)

As to claims 5, 10, 17, 24, 29, and 34, Nanjo et al as modified teaches wherein the calculated values are weighted based on a ranking defined by relevance of the identified documents, such that substrings that occur in more relevant ones of the identified documents are assigned higher calculated values than substrings that occur in less relevant ones of the documents (see Pant et al, column 7, lines 23-30, and see column 10, lines 36-49.)

As to claim 6, Nanjo et al teaches a method of locating documents (see column 4, lines 4-18) in response to a search query (see column 6, lines 22-26), the method comprising:

receiving the search query from a user (see column 6, lines 22-26);
generating a list of relevant documents based on search terms of the query (see column 6, lines 30-35);

identifying a subset of documents that are most relevant ones of the documents in the list of relevant documents (see column 17, lines 35-47, where “most relevant document” is read on “exact matches”.)

generating a plurality of multiword substrings from the query (see column 4, lines 4-12, and see column 8, lines 21-33, where “multiword substrings” is read on “combination of text strings”) in which each of the substrings includes at least two words (see column 8, line 33. It is inherent that a “phrase” contains at least two words”);

refining the generated list of relevant documents based on the selected semantic units (see column 19, lines 15-25, where “refining” is read on “optimizing”).

Nanjo et al does not teach:

calculating, for each of the generated substrings, a value related to one or more documents in the subset of documents that contain the substrings; and selecting semantic units from the generated multiword substrings based on the calculated values.

Pant et al teaches a system for relevance-ranking of search results (see Abstract), in which he teaches calculating, for each of the generated substrings, a value related to one or more documents in the subset of documents that contain the substrings (see column 7, lines 7-50); and selecting semantic units from the generated multiword substrings based on the calculated values (see column 10, lines 38-49, and see column 13, lines 9-25.)

Therefore, it would have been obvious to a person having ordinary skill in the art at the time the invention was made to have modified Nanjo et al as modified to include calculating, for each of the generated substrings, a value related to one or more documents in the subset of documents that contain the substrings; and selecting semantic units from the generated multiword substrings based on the calculated values.

It would have been obvious to a person having ordinary skill in the art at the time the invention was made to have modified Nanjo et al as modified, by the teaching of Pant et al, because calculating, for each of the generated substrings, a value related to one or more documents in the subset of documents that contain the substrings; and selecting semantic units from the generated multiword substrings based on the calculated values, would enable the system to define and apply relevancy values (factors) between documents as well as

between contents of documents in order to identify the most relevant (strongest) match between the search terms and the identified terms within a query.

As to claims 7 and 31, Nanjo et al as modified teaches wherein the identified subset includes a predetermined number of the most relevant ones of the documents in the list of relevant documents (see Nanjo et al, column 17, lines 35-50, where “most relevant” is read on “exact match”.)

As to claim 11, Nanjo et al teaches a system (see Abstract) comprising:
a server connected to a network (see column 12, lines 7-12, where in a “networked environment”, the “computer system 400” plays the role of “a server”), the server receiving search queries from users via the network (see column 6, lines 22-26), the server including:
at least one processor (see column 11, lines 28-30, and see column 12, lines 7-12); and
a memory operatively coupled to the processor (see column 11, line 30), the memory storing program instructions that when executed by the processor (see column 11, lines 57-67), cause the processor to:

identify a list of documents (see column 4, lines 4-18) relating to the search query (see column 6, lines 22-26, and see lines 30-35) by matching individual search terms in the query to an index of a corpus (see Abstract, and see column 21, lines 29-30); a plurality of multiword substrings from the query (see column 4, lines 4-12, and see column 8, lines 21-33, where “multiword substrings” is read on “combination of text strings”) in which each of

the substrings includes at least two words (see column 8, line 33. It is inherent that a “phrase” contains at least two words”.)

For the teachings of: calculate, for each of the generated substrings, a value relating to one or more documents of the identified list of documents that contain the generated substrings; and select semantic units from the generated multiword substrings based on the calculated values, applicants are kindly directed to remarks and discussions made in claims 1 and 6 above.

As to claim 12, Nanjo et al as modified teaches wherein the processor refines the identified list of documents based on the selected semantic units (see Nanjo et al, column 19, lines 15-25, where “refining” is read on “optimizing”.)

As to claims 13 and 20, Nanjo et al as modified teaches wherein the system transmits the refined list of documents to the user (see Nanjo et al, column 3, lines 19-28, where “transmitting to the user” is read on “displaying to the user.”)

As to claim 14, Nanjo et al as modified teaches a network (see Nanjo et al, column 12, lines 7-12.)

Nanjo et al as modified still does not teach wherein the network is the Internet and the corpus is a collection of web documents.

Pant et al, in another embodiment of his invention, teaches wherein the network is the Internet (see column 3, lines 22-32) and the corpus is a collection of web documents (see

Art Unit: 2175

column 2, lines 4-10, where “web documents” is read on “documents written in SGML, HTML, or other mark-up languages”, and see column 5, lines 61-63.)

Therefore, it would have been obvious to a person having ordinary skill in the art at the time the invention was made to have modified Nanjo et al as modified to include wherein the network is the Internet and the corpus is a collection of web documents.

It would have been obvious to a person having ordinary skill in the art at the time the invention was made to have modified Nanjo et al as modified, by the further teaching of Pant et al, because wherein the network is the Internet and the corpus is a collection of web documents, would expand the usability of the system across the universe by connecting to the internet and searching web-based documents from any remote location accessible by a computer to the network.

As to claim 18, Nanjo et al teaches a server (see column 12, lines 7-12, where in a “networked environment”, the “computer system 400” plays the role of “a server”) comprising:

a processor (see column 11, lines 28-30, and see column 12, lines 7-12); and
a memory operatively coupled to the processor (see column 11, line 30), the memory including:

a ranking component (see column 26, lines 31-35) configured to return a list of documents (see column 4, lines 4-18) ordered by relevance in response to a search query (see column 6, lines 22-26, and see lines 30-35); and

Art Unit: 2175

a semantic unit locator component configured to locate semantic units (see column 10, lines 1-4), each having a plurality of words (see column 8, line 33. It is inherent that a “phrase” contains a plurality of words”), in search queries entered by a user (see column 7, lines 15-18.)

For the teaching of: locating semantic units based on a predetermined number of most relevant documents in the list of documents returned by the ranking component, applicants are kindly directed to remarks and discussions made in claims 1 and 6 above.

As to claim 19, Nanjo et al as modified teaches the server further including: a search engine (see Nanjo et al, column 15, lines 31-38) configured to refine the list of documents based on the located semantic units (see Nanjo et al, column 19, lines 15-25, where “refining” is read on “optimizing”.)

As to claim 21, applicants are kindly directed to the remarks and discussions made in claims 1, 6, and 18 above.

As to claim 25, Nanjo et al teaches a computer-readable medium storing instructions for causing at least one processor to perform a method that identifies semantic units within a search query (see Abstract, and see column 11, line 28 through column 12, line 12.)

For the remaining teachings of this claim, applicants are directed to the remarks and discussions made in claims 1 and 6 above.

As to claim 30, applicant is directed to the remarks and discussions made in claims 1, 6, 11, 18, and 25 above.

As to claim 36, applicant is directed to the remarks and discussions made in claims 1, 6, 11, 18, 25, and 30 above.

As to claims 37, 38, 39, 40, and 41, Nanjo et al as modified teaches wherein the calculated values are weighted based on a ranking defined by relevance of the identified documents, such that an occurrence of a substrings in a more relevant one of the documents is weighted more than an occurrence of the substrings in a less relevant one of the documents (see Pant et al, column 7, lines 7-50, and see column 10, lines 38-49.)

4. Claims 4, 9, 16, 23, 28, and 33 are rejected under 35 U.S.C. 103(a) as being unpatentable over Nanjo et al (U.S. patent No. 5,778,361) in view of Pant et al (U.S. Patent No. 6,012,053), as applied to claims 1-3, 5-8, 10-15, 17-22, 24-27, 29-32, 34, and 36-41 above, and further in view of Tso et al (U.S. Patent No. 6,385,602.)

As to claim 4, 9, 16, 23, 28, and 33, Nanjo et al as modified teaches wherein the selection of the semantic units further includes ranking search results based on a relevance factor (see Pant et al, column 7, lines 15-30, where “more relevant” and “less relevant” terms are taught.)

Nanjo et al as modified still does not teach discarding the generated substrings that overlap other ones of the generated substrings with higher calculated values.

Art Unit: 2175

Tso et al teaches dynamic categorization of search results (see Abstract), in which he teaches discarding the generated substrings that overlap other ones of the generated substrings with higher calculated values (see column 5, lines 9-28.)

Therefore, it would have been obvious to a person having ordinary skill in the art at the time the invention was made to have modified Nanjo et al as modified, to include discarding the generated substrings that overlap other ones of the generated substrings with higher calculated values.

It would have been obvious to a person having ordinary skill in the art at the time the invention was made to have modified Nanjo et al as modified, by the teaching of Tso et al, because discarding the generated substrings that overlap other ones of the generated substrings with higher calculated values, would enable the system to present the most relevant results to the user in a search query by discarding the hits with lower relevance factors. For example, a search criteria that results in twenty thousand hits, discards nineteen thousand nine hundred less relevant hits and only displays the one hundred highest relevant hits to the user, as taught by Tso et al (see column 5, lines 28-32.)

5. Claim 35 is rejected under 35 U.S.C. 103(a) as being unpatentable over Nanjo et al (U.S. patent No. 5,778,361) in view of Pant et al (U.S. Patent No. 6,012,053), as applied to claims 1-3, 5-8, 10-15, 17-22, 24-27, 29-32, 34, and 36-41 above, and further in view of Freimann et al (U.S. Patent No. 6,134,554.)

Art Unit: 2175

As to claim 35, Nanjo et al as modified still does not teach wherein the computer-readable medium is a CD-ROM, floppy disk, tape, flash memory, system memory, hard drive, or data signal embodied in a carrier wave.

Freimann et al teaches a system and method for recording receipt of information (see Abstract), in which he teaches wherein the computer-readable medium is a CD-ROM, floppy disk, tape, flash memory, system memory, hard drive, or data signal embodied in a carrier wave (see column 14, lines 17-21.)

Therefore, it would have been obvious to a person with ordinary skill in the art at the time the invention was made to have modified Nanjo et al as modified to include wherein the computer-readable medium is a CD-ROM, floppy disk, tape, flash memory, system memory, hard drive, or data signal embodied in a carrier wave.

It would have been obvious to a person having ordinary skill in the art at the time the invention was made to have modified Nanjo et al as modified by the teaching of Freimann et al, because the computer-readable medium being a CD-ROM, floppy disk, tape, flash memory, system memory, hard drive, or data signal embodied in a carrier wave, would increase the system flexibility to incorporate any/all means of data storage in order to maintain maximum flexibility and compatibility with a variety of hardware systems featuring any of the mentioned variety of storage means.

Response to Arguments

6. Applicant's arguments filed on 23-March-2004 with respect to the rejected claims in view of the cited references have been fully considered but they are not deemed persuasive:

In response to the applicants' arguments that "the term 'semantic unit', as defined by the pending application, refers to multiple terms that are considered to function as a 'compound' that forms a single semantically meaningful unit. (Spec., page 2)", the arguments have been fully considered but are not deemed persuasive, because "multiple terms that are considered to function as a 'compound' that forms a single semantically meaningful unit" is not recited in the rejected claim.

Further, in response to the applicant's arguments that "nothing in Nanjo discloses or suggests that these sub-strings correspond to multi-word substrings that form a single semantic unit (i.e., a semantically meaningful unit)", the arguments have been fully considered but are not found persuasive because "sub-strings correspond to multi-word substrings that form a single semantic unit (i.e., a semantically meaningful unit)" is not recited in the rejected claim.

In view of the above arguments, the examiner notes that although the claims are interpreted in light of the specification, limitations from the specification are not read into the claims. See *In re Van Geuns*, 988 F.2d 1181, 26 USPQ2d 1057 (Fed. Cir. 1993).

In response to the applicants' arguments that "although Pant may generally calculate 'values' based on the terms in a document and search queries, this does not even suggest

calculation of values that are then used to select semantic units as recited in claim 1”, the arguments have been fully considered but are not deemed persuasive, because Pant et al teaches “relevance factor” being the “order of search terms”, which means that “semantic units” can be searched (“selected”) by the calculated values (“relevance factor”.) See column 13, lines 17-25.)

In response to the applicant’s argument regarding the dependent claims 4, 9, 16, 23, 28, and 33, the argument has been fully considered but is not found persuasive in view of the remarks and discussions made in response to applicants’ arguments above.

Conclusion

7. **THIS ACTION IS MADE FINAL.** Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE MONTHS from the mailing date of this action. In the event a first reply is filed within TWO MONTHS of the mailing date of this final action and the advisory action is not mailed until after the end of the THREE-MONTH shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the mailing date of this final action.

Art Unit: 2175

8. Any inquiries concerning this communication or earlier communications from the examiner should be directed to Tony Mahmoudi whose telephone number is (703) 305-4887. The examiner can normally be reached on Mondays-Fridays from 08:00 am to 04:30 pm.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Dov Popovici, can be reached at (703) 305-3830.

tm

May 18, 2004



DOV POPOVICI
SUPERVISORY PATENT EXAMINER
TECHNOLOGY CENTER 2100